DOCUMENT RESUME

ED 124 897 CS 002 741

| | |
|---|---|
| AUTHOR | Sullivan, Howard J.; Niedermeyer, Fred C. |
| TITLE | Pupil Achievement under Varying Levels of Teacher Accountability. Professional Paper 28. |
| INSTITUTION | Southwest Regional Laboratory for Educational Research and Development, Los Alamitos, Calif. |
| SPONS AGENCY | Office of Education (DHEW), Washington, D.C. |
| REPORT NO | SWRL-TN-5-70-37 |
| PUB DATE | 73 |
| NOTE | 11p. |
| EDRS PRICE | MF-$0.83 HC-$1.67 Plus Postage. |
| DESCRIPTORS | *Educational Accountability; Elementary Education; *Reading Achievement; *Reading Programs; Reading Research |

ABSTRACT

This year-long study was conducted to investigate effects of various teacher accountability factors on reading achievement of first-grade pupils. Each of fifteen schools was assigned through randomized blocking procedures (blocking factors were reading achievement and percentage of minority students) to one of four levels of accountability. Accountability levels were defined from low to high as follows: use of the regular first-grade reading program without supplementation; the regular program plus a supplementary system including instructional objectives, regular assessment and remedial exercises, and objectives-keyed pupil performance records for each of the fifteen instructional units in the program; the second condition plus either a pacing schedule or reporting of unit scores to the school principal; and the second condition plus both pacing and reporting. A 45-item criterion test based on the major objectives of the reading program was individually administered at the end of the school year. An L-test for four rankings (accountability levels) by three observations (mean subtest scores on three major reading objectives) revealed a significant positive relationship ($p < .001$) between level of teacher accountability and pupil reading achievement. (Author)

**********
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
**********

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED124897

SWRL EDUCATIONAL RESEARCH AND DEVELOPMENT

Professional Paper 28

December 1973

# PUPIL ACHIEVEMENT UNDER VARYING LEVELS OF TEACHER ACCOUNTABILITY

Howard J. Sullivan and Fred C. Niedermeyer

ABSTRACT

This year-long study was conducted to investigate effects of various teacher accountability factors on reading achievement of first-grade pupils. At the beginning of the school year, each of 15 schools was assigned through randomized blocking procedures (blocking factors were reading achievement and percentage of minority students) to one of four levels of accountability. Accountability levels were defined from low to high as follows: (1) use of the regular first-grade reading program without supplementation, (2) the regular program plus a supplementary system including instructional objectives, regular assessment and remedial exercises, and objectives-keyed pupil performance records for each of the 15 instructional units in the program, (3) condition 2 above plus either a pacing schedule or reporting of unit scores to the school principal, (4) condition 2 plus both pacing and reporting. A 45-item criterion test based on the major objectives of the reading program was individually administered at the end of the school year. An L-test for four rankings (accountability levels) by three observations (mean subtest scores on three major reading objectives) revealed a significant positive relationship ($p < .001$) between level of teacher accountability and pupil reading achievement.

CS 002 741

PUPIL ACHIEVEMENT UNDER VARYING LEVELS OF TEACHER ACCOUNTABILITY

Howard J. Sullivan and Fred C. Niedermeyer[1]

The idea of teacher accountability for producing demonstrated gains in pupil achievement currently enjoys considerable popularity among educators. In general, however, specific procedures for promoting accountability have neither been well defined and tested nor installed in the schools. Little, if any, experimental research has been conducted to investigate the effects of accountability procedures on pupil achievement.

The present study involved an attempt to (1) identify and operationalize a set of factors related to accountability and (2) empirically determine the relationship between these factors and learner achievement. Accountability factors identified and investigated in the study were as follows:

- Specification of what the teacher is accountable for--that is, what objectives must be attained--and provision of a means for regular determination by the teacher of whether the objectives are being attained.
- Development of a schedule indicating to the teacher the maximum amount of time in which pupils should attain the objectives for which the teacher is to be held accountable.
- Requirement of regular reporting of pupil attainment of objectives to a person (in this case, the person was the school principal) to whom the teacher is accountable.

Each accountability factor was operationalized using materials and procedures that could be developed cooperatively by teachers and other individuals with responsibilities for the content and outcomes of instruction at a district, county, or state level.

---

[1]The authors wish to express their appreciation to Mrs. Marjorie Craig of the San Diego Unified School District for her assistance throughout the study.

The hypothesis investigated in this study was that there is a positive relationship between level of teacher accountability and learner achievement on the instructional objectives for which the teacher is held accountable. The dependent variable was end-of-year reading achievement of first-grade pupils on the inferred instructional objectives of the regular Grade 1 reading program in the school district in which the study was conducted. The lowest of the four levels of teacher accountability in the study involved use of the regular reading program without instructional objectives or any supplementary information related to accountability. For this level, no special accountability factors were included. The highest level of accountability involved combined use of all the following factors: the regular reading program accompanied by a set of instructional objectives for the program, tests for regular assessment of learner attainment of the objectives, a pacing schedule indicating the number of weeks in which given sets of objectives should be mastered, and the requirement that teachers submit to their principals a record of pupil achievement scores on each regular assessment exercise administered throughout the year.

## METHOD

### Subjects

The Ss consisted of all first-grade pupils in 15 schools in a metropolitan southern California school district. The 15 schools were selected from among 60 schools in the district on the basis of their similarity on matching factors employed in the study and described below. The total sample included 1889 children from 67 first-grade classes.

### Materials

The basic instructional program used with all groups in the study was a commercially published Grade 1 reading program that is both the most widely used of the three California state-adopted primary-grade reading programs and the basic program used in the district in which the study was conducted. The complete Grade 1 program includes over 600 new words, plus content in areas related to word decoding, comprehension, and study skills. The program materials contain neither an explicit statement of the instructional objectives of the program nor a concise summary of the content and skills that children are expected to learn.

The basis for the accountability procedures used in three experimental treatments was a set of objectives-based assessment and remediation materials designed to supplement the basic instructional program in reading. This supplementary system was designed to enable the teachers to (1) identify explicitly the skills and related content that children should learn from the basic program, (2) determine at regular assessment intervals whether the intended learning has occurred, and (3) provide remedial instruction on each skill that had not been mastered by the conclusion of an assessment period.

The initial steps in design of the supplementary system involved inferring of the instructional objectives of the basic program and subsequent identification of the inferred objectives judged to be important enough that teachers should be held responsible for their attainment. From among a total of seven objectives inferred for the program, three were judged to be critical to the development of reading skills for first-grade children. These three objectives were the ability to (1) read the new basic words in the program, (2) say the sounds of various letters and letter combinations contained in the program and (3) sound out and read new words composed of the letters and letter combinations.

As a basis for indicating clearly to the teacher the objectives that children should attain and for providing for regular assessment and remediation, the objectives and content for the basic program were divided into 15 units containing relatively equal amounts of new content to be attained by the pupils. For each of the 15 units, the supplementary program contained a listing in the teachers' manual of all content (i.e., new basic words, sounds, and decoding words) for each objective, a simple teacher-administered group test measuring pupil performance on each objective, a Class Record Sheet for recording each pupil's achievement by objective on each test, and brief remediation exercises keyed to each objective and designed for use with pupils not attaining a specified criterion score on one or more unit objectives. Thus, a classroom set of supplementary materials consisted of a teachers' manual for the supplementary system, 15 tests (one per unit) for each child, 15 Class Record Sheets, and 15 sets of short remedial exercises for each child to use only if his unit-test performance indicated a need for remediation on one or more objectives. The total cost of the year-long supplementary system, which was designed to be both simple and inexpensive, was less than 35 cents per pupil.

Procedures

Four levels of teacher accountability were designed as the experimental treatments. These levels of accountability are described in order below, from lowest to highest level.

Level 1: Program Only. Teachers at the schools under this condition used only the state-adopted commercial reading program serving as the basic instructional program in the study. Since this group of teachers used their regular reading program without supplementation or experimental intervention other than posttesting, it seems safe to assume that they taught reading as they normally do.

Level 2: Program plus Objectives-based Assessment System (OAS). This condition involved use of the basic instructional program plus the supplementary system designed for use with it. Thus, in addition to the basic instructional program, teachers in this group were provided with a unit-by-unit listing of the objectives that their pupils should attain, tests to determine whether pupils had attained the objectives, and short remedial exercises keyed to the unit objectives.

Level 3: Program plus OAS plus Pacing OR Reporting. This condition included use of both the basic instructional program and the objectives-based supplementary system, plus one additional accountability factor. For teachers in three of the six schools in this treatment, the added accountability factor was a pacing schedule specifying for each of the 15 units the maximum number of weeks that should be required for pupils to attain all objectives of the unit. The pacing schedule for the year was developed to permit completion of the entire Grade 1 program. Instead of the pacing schedule, the added accountability factor for teachers in the other three schools in this treatment was a reporting requirement. Following each unit test, each teacher under the reporting condition was required to submit to her principal a copy of the Class Record Sheet showing the test score of each pupil on each objective for the unit. (This level was considered to fall between Levels 2 and 4 because it included one additional factor--either pacing or reporting--to Level 2, but not both additional factors in combination, as in Level 4. Including of the pacing and reporting factors separately under this condition permitted analysis of their individual effects.)

Level 4: Program plus OAS Pacing AND Reporting. This condition involved use of the basic instructional program with all three accountability factors: the objectives-based supplementary system, the pacing schedule, and the reporting requirement. The Class Record Sheets used for recording unit test scores and reporting them to principals for this group also contained the recommended pacing schedules for each unit, thus enabling the principal to quickly determine whether the rate of progress of a group or class was consistent with the recommended rate.

The school served as the experimental unit in the study. From among the 60 available schools in the district in which the study was conducted, five schools each were selected to represent low, middle and high achievement levels in reading. The particular schools at each achievement level were selected because of their similarity on (1) end-of-first-grade mean scores from the preceding school year on the California state-mandated Cooperative Reading Test and (2) percentage of minority group students. Randomized blocking procedures were employed to assign the 15 schools to treatments. From within each block of five schools, two schools were assigned to the Program plus OAS plus Pacing or Reporting condition (one school to Pacing, one to Reporting) and one school each to the other three levels of accountability. Thus, a total of three schools, one from each of the three levels of the blocking factor, were assigned to each treatment except the Pacing or Reporting treatment. Two schools at each blocking level, or a total of six schools were assigned to the Pacing or Reporting treatment to permit analysis of possible differential effects of the pacing and reporting requirements.

A 90-minute teacher-training session for each experimental treatment (Levels 2, 3 and 4) was held at the beginning of the school year to train each group of teachers in the procedures that they were to use. The training sessions were arranged and conducted by a primary-grade supervisor from the participating district, using materials and procedures provided by the experimenters. Because accountability procedures may be threatening to many teachers, a concentrated effort was made to describe the supplementary system and requirements in as pleasant and non-threatening a manner as possible, both in the written materials and in the training session.

Data Sources

A 45-item oral-response criterion test was developed to assess end-of-year achievement on the three instructional objectives of the reading program (1 - reading program words, 2 - saying sounds of letters and letter combinations, and 3 - reading new words composed of letters and letter combinations from the program). One item per objective for each of the first 11 units of the program was randomly selected from the population of all available items and included in the test. Since Units 12-15 of the program included content only for the "reading program words" objective, the three randomly selected items for each of these four units assessed performance only on the program-words objective. Thus, the test contained 23 items assessing performance on this objective and 11 items each (one item for each of the first 11 units) for the letter-sounds and word-decoding objectives.

The criterion test was individually administered by trained examiners during the next-to-last week of the school year to a sample of 24 children per school selected at random from among all first graders in each of the 15 schools. The 360 posttested children, therefore, included a total of 144 from the six Pacing or Reporting schools (72 for each condition) and a total of 72 children from each group of three schools at each of the three remaining accountability levels. The KR-21 reliability coefficient, computed on the sample of 360 posttested Ss, was .96 for the complete 45-item posttest. KR-21 coefficients for the subtests assessing performance on the three objectives ranged from .92 to .95.

Data were also collected on the amount of the program completed at each accountability level. At the time of posttesting, teachers were asked to indicate the number of the unit in the reading program in which each tested child was working. Each teacher was also interviewed briefly at this time to determine whether any contamination of experimental procedures had occurred across accountability levels.

Data Analysis

Posttest mean scores were tabulated for each treatment group on each objective. An L-test (Page, 1963) for four rankings (accountability levels) by three observations (subtest scores on each of the

three objectives) was performed to determine the relationship between level of teacher accountability and end-of-year pupil achievement on the instructional objectives of the reading program.

## RESULTS

The posttest mean scores by treatment and objective are shown in Table 1. Analysis of the table reveals that, without exception, pupil achievement increased on each objective with each increase in the level of teacher accountability. Mean percentage scores on the complete posttest ranged from 55 percent for the program-only group representing the lowest level of accountability in the study to 68 percent for Ss at the highest level of accountability--the Pacing and Reporting group. The L-test for four accountability levels by three subtest observations revealed a significant monotonic relationship ($L = 90$, $p < .001$) between level of teacher accountability and pupil achievement on the instructional objectives of the program.

Table 1

Posttest Mean Scores by Accountability Level and Objective

| Accountability Level | Objective | | | Total Test (45) |
|---|---|---|---|---|
| | Words (23 items) | Letter Sounds (11 items) | Word-Decoding (11 items) | |
| 1 - Program only (Lowest level) | 13.3<br>58% | 6.4<br>58% | 5.1<br>46% | 24.8<br>55% |
| 2 - Program + Objectives-based Assessment System (OAS) | 14.1<br>61% | 6.9<br>63% | 6.0<br>55% | 27.0<br>60% |
| 3 - Program + OAS + Pacing or Reporting | 14.9<br>65% | 7.2<br>65% | 6.4<br>58% | 28.5<br>63% |
| 4 - Program + OAS + Pacing & Reporting | 16.1<br>70% | 7.9<br>72% | 6.7<br>61% | 30.7<br>68% |

Note: $L = 90$, $p < .001$ for 4 rankings x 3 observations.

Mean scores were also computed and compared for the two sub-groups within the Program plus OAS plus Pacing or Reporting treatment. The difference in mean score on the total test between Ss under the pacing condition and Ss under the reporting condition within this treatment was less than one raw-score point (Pacing $\bar{x}$ = 28.2, Reporting $\bar{x}$ = 28.2).

The mean number of units completed out of the total number of 15 units in the reading program is shown below by accountability level:

| Accountability Level | Units Completed |
|---|---|
| 1 - Program only | 9.0 (60% of total program) |
| 2 - Program + OAS | 10.9 (73%) |
| 3a - Program + OAS + Reporting | 9.7 (65%) |
| 3b - Program + OAS + Pacing[3] | 10.8 (72%) |
| 4 - Program + OAS + Pacing and Reporting | 11.2 (75%) |

As shown above, the program completion rate ranged from 60% at the lowest accountability level to 75% at the highest level. Conditions which included a pacing schedule (Levels 3b and 4) generally showed a higher completion rate than the other levels. However, the general trend toward decreasing completion rates with a decrease in accountability level was broken by the high completion rate at Level 2. The high completion rate at this level was due to the fact that subjects in the low-achievement school in this group completed an average of 11.6 units, the second highest number of units completed by any school in the study and more than two units higher than the average number completed by the other low-achievement schools. Interviews with the teachers at this school revealed that their principal had obtained and established the pacing schedule used at Levels 3 and 4 and was collecting reporting sheets on a regular basis. Thus, this one Level 2 school was essentially under the Level 4 Pacing and Reporting condition.

## DISCUSSION

This study was conducted to investigate the relationship between level of teacher accountability and the reading achievement of first-grade pupils. The data revealed a significant monotonic relationship between accountability level and pupil achievement. An increment of end-of-year pupil achievement on the instructional objectives of the first-grade reading program used in the study occurred with each increase in level of teacher accountability.

---

[3]The two conditions under Level 3 are shown separately in this listing to indicate the difference in program completion rate associated with the pacing condition as contrasted with the reporting condition.

One important aspect of the accountability factors employed in the study was the provision for frequent en route assessment and reporting of learner performance on specific instructional objectives. Often, pupil achievement data are systematically collected within a school or district only through use of standardized tests administered at the beginning and/or end of the school year. Norm-referenced data collected at such infrequent intervals are of relatively little use to the teacher in making instructional decisions that will increase the chances for attainment of the objectives for which the teacher is held accountable.

Because of the type of information sought from the experiment, the potential value of the objectives-based en route assessment and reporting procedures were not fully exploited in the present study. For experimental purposes, the reporting condition was intentionally limited to regular collection of the reports in order to permit analysis of the effects of reporting per se, and no assistance or further directions designed to yield improved performance were provided for teachers who submitted Class Record Sheets indicating sub-standard pupil achievement and/or pacing. However, in actual operation of an effective accountability system, the regular reports can be used by supervisory personnel at the school or district level to identify teachers who need help in improving their instruction. Such a formative evaluation procedure used in conjunction with a teacher accountability system should result in improved instruction and concomitant improvements in pupil achievement.

The accountability factors identified and operationalized as part of the present study included provision of (1) learner objectives and a regular means for assessing their attainment, (2) a pacing schedule to encourage and help teachers plan for an acceptable rate of progress through the instructional program, and (3) a procedure for regular reporting of en route learner achievement to school officials. Data from the study reveal a significant positive relationship between the presence of these factors and pupil achievement in first-grade reading. It seems probable that, in the operation of a well-designed accountability system in the schools, the positive effects of the accountability factors from this study could be increased beyond the levels obtained in the study by incorporating intervention procedures designed to improve the performance of teachers having sub-standard pacing and/or pupil achievement. Potentially then, application in the schools of the accountability factors identified and employed in this study should yield positive results both in implementing an effective teacher-accountability system and in improving pupil achievement.

## REFERENCE

Page, E.B. Ordered hypotheses for multiple treatments: A significant test for linear ranks. In The 20th Yearbook of the National Council on Measurement in Education, 1963, pp. 17-23.